Raj Singh
Syncline, Inc.
PO Box 390862
Cambridge, MA 02139
phone: 617.661.7134
fax: 617.661.7135
email: rs@syncline.com

PROVIDING GIS SERVICES OVER THE INTERNET

Abstract: The basic theme of this presentation is to discuss the organizational impacts of the Internet on the way we provide GIS services intra-agency and to the public.

The Internet and the Web are highly useful tools in our GIS arsenal, but we must get beyond the hype and understand what they truly offer. For our purposes, the Internet is most important as a network implementation tool allowing programs on different computers to talk to each other—a "client-server environment." This enables us to separate the GIS user interface from the actual GIS server (containing data and the spatial analytical processes) in many ways. The Web is important because it provides a user interface that most everyone already understands, namely a browser. Placing our GIS clients in this Web-based interface allows us to leverage the widespread knowledge of how to use a web site, pre-existing security and authentication mechanisms, low (or zero) software costs on the client end, and tremendous scalability.

There are two main hurdles to realizing these web-based benefits, technology development and organizational creativity. The technology, albeit new, is in place today for displaying geography over the Web. The spectrum spans simple GIFs and JPEGs of maps and imagemaps, to more robust and sophisticated implementations using ESRI's ArcView Internet Map Server, MapInfo's MapXtreme, Live Picture's Image Server and LizardTech's MrSID Image Server. In examining the state of the technology, Web-based Mapping is divided into three separate tiers: Map Viewing, Web Mapping, and Web-based GIS.

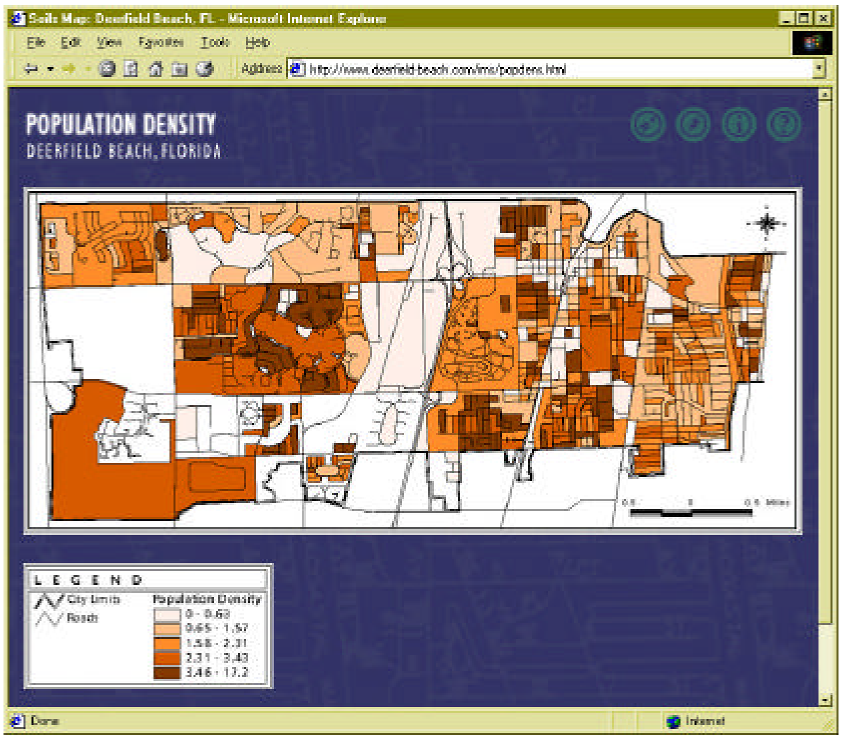
While many of the technical tools exist today, many organizations have yet to completely grasp the realm of possibilities afforded by these technologies. A discussion of several different scenarios, including outsourcing the maintenance of GIS servers and creating self-service mapping centers on the Web for the public, will shed light on the true promise of Web-based GIS.

ANALYSIS OF WEB-BASED MAPPING TECHNOLOGIES

Tier 1: Map Viewing

A simple, effective way of distributing geographic information is to save maps in an HTML-compatible format and post them on a Web site. This, of course, limits the size, color, and resolution of the map to a standard computer monitor whose base specifications are 800 x 600 pixels at 72 dots per inch and displaying about 200 colors (the "web-safe" color palette). This method may work well for simple maps that rarely change, such as thematic maps of the Census' demographic data (Figure 1). However, this method is labor intensive and does not scale well, thus only a few maps are likely to be posted on the Web site. The implication is that the maps generated are "canned" and therefore must have universal appeal. Many people may find the "canned" maps inadequate, particularly for those who need more detailed or customized information. In sum, these maps are not interactive and do not take advantage of the Internet's most compelling attributes.

FIGURE 1: AN HTML-FORMATTED MAP OF POPULATION DENSITY



Tier 2: Web Mapping

The next technology tier addresses many of the aforementioned shortcomings including the amount of data able to be shown and the level of detail issue. This tier also allows the map to be linked to other Web-accessible files. There are many tools that allow the end user to zoom and pan a map. The simplest way to gain functionality is to make your map an HTML imagemap. All common web browsers can read the HTML code that links areas of the image to other Web-accessible files. This technique was used in one project to link a map of the harbor to photographs and sections of a report (http://www.syncline.com/rs/nb/harbor.html).

The next step up is to take advantage of a product called Acrobat from Adobe (http://www.adobe.com). Most people are familiar with the PDF file format and the Acrobat Reader, which is a free program that allows you to view PDF files, and most people know that you can zoom and pan PDF files. However, many people do not know that PDF files can include links, much like an imagemap. This can be a powerful tool for disseminating static maps, as it requires no new server software, and is inexpensive and easy to use compared to the more advanced applications that follow.

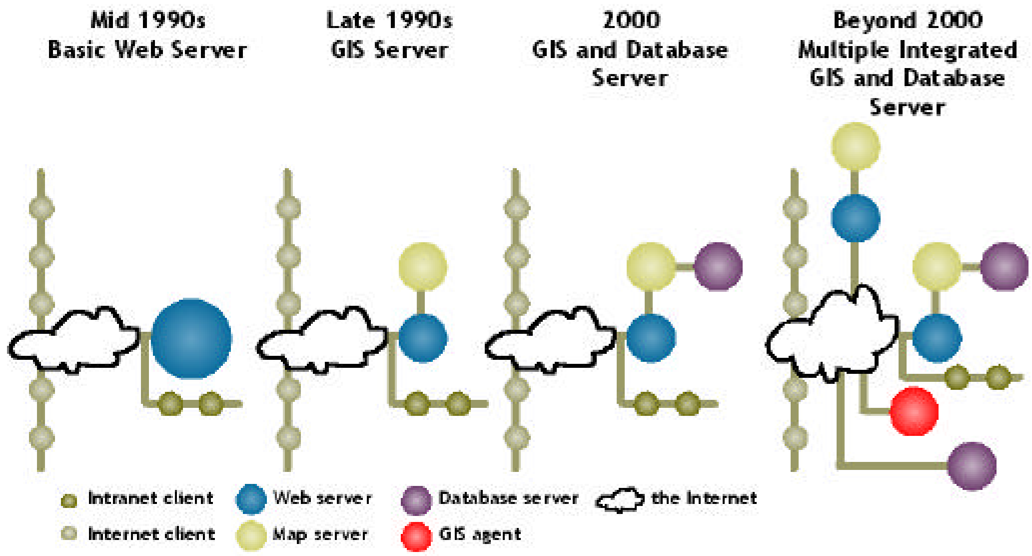
Syncline used LivePicture's (http://www.livepicture.com) Flashpix image server to Web-enable a 100Mb orthophoto of Deerfield Beach (http://www.deerfield-beach.com/html/citytour.html). The image server is able to resample the original image on the fly to send a small file to the client's Web browser. It is also able to clip small, high-resolution pieces out of the original image, allowing the user to zoom in and gain resolution. LivePicture's Web client applications allow areas of the image to be hotlinked to other Web-accessible files. Syncline used this technology to create a "virtual city tour" of Deerfield Beach ( ). The orthophoto is linked to pictures, movies and panoramic images. Another similar technology is MrSID from LizardTech (http://www.lizardtech.com). The main differences between the two products right now is that MrSID allows you to save a smaller file on the server, but it doesn't support hotlinking to other files and the LivePicture image server actually streams the image to the client, resulting in better performance. If the image is also used in a GIS, MrSID has the added advantage of storing geographic coordinate information and is supported by ESRI's ArcView GIS (both products' file formats can be used in Adobe Photoshop).

FIGURE 2: DEERFIELD BEACH VIRTUAL CITY TOUR

Tier 3: Web-based GIS (WGIS)

The technologies described up to this point are not suitable for situations where users need access to constantly changing information, or where each user's mapping needs are unique. Meeting these demands requires a true "Internet-enabled" GIS system. The first products made available by companies such as ESRI and MapInfo focussed on dynamic mapmaking (Figure 3, "Late 1990s"). They read the data formats of their desktop GIS programs, created a GIF or JPEG map based on the user's chosen data layers and zoom level and displayed this image in an HTML page or a Java applet. The current state of the art in WGIS is maintaining attribute and/or geometry data in a relational database management system (RDBMS) (Figure 3, "2000"). This can improve data retrieval performance and makes it possible to take advantage of the capabilities of RDBMS technologies, such as access control, parallelization, portability, security, concurrency, replication, distributed database, query and client/server architectures.

FIGURE 3: WGIS NETWORK INFRASTRUCTURE



A NETWORK-CENTRIC GIS ARCHITECTURE

While corporate applications will be most concerned with exploiting these systems for scalability and performance, I believe that the network-centric nature of the systems architecture offers great opportunity to public sector efforts to disseminate geographic information internally and to the public. Current WGIS and RDBMS technology allows servers to be distributed all across the Internet while still working together. There is also the possibility that GIS systems from different vendors will work together using interoperability standards such as OpenGIS's Simple Features Specification for SQL (http://www.opengis.org/techno/specs.htm). These trends taken together create a very different GIS environment than that which exists today (Figure 3, "Beyond 2000") and raises many technological and organizational issues regarding security, privacy, quality management, public relations, network management, compatibility, training, cost control, etc. What is the place for the Internet in our organizations? How do we manage such an open technology?

I will conclude this paper with a description of a scenario in which a state integrates WGIS services from the town level up to the state. As a point of reference I will describe a framework being contemplated by the Essex County, Massachusetts Registry of Deeds in conjunction with MassGIS, the State's geographic data clearinghouse. Like many states, Massachusetts has a well-established state-level GIS agency, MassGIS, that has built up a large library of data. As MassGIS is a part of the Department of Environmental Affairs, their database is largely comprised of natural resource information, but also includes base data such as orthophotographs, Census demographics, political boundaries and infrastructure. Towns have varying levels of GIS sophistication, but there is a desire to create some standard level of Internet-based GIS services across the state or at least the county. For example, a news article early in 1999, stated the Secretary of State's interest in having a statewide zoning database that would allow potential developers to identify all locations in the state that met their siting requirements for a new project. Essex County is interested in providing real time access to parcel ownership information on the Internet to reduce the need for people to come into their main office every day.

The most straightforward solution to these needs is to have one central GIS that is updated remotely by the towns. Good tools already exist to allow relational databases to be updated through forms in a web browser. Syncline is developing a Web-based Java application that would allow a user to update geographic features. So technologically this solution is possible, but it is not practical from an organizational standpoint because municipalities are not likely to surrender control of their information systems to a state authority.

The most likely scenario is a hybrid system in which some of the more advanced towns manage their own Internet-aware GIS, while the smaller towns that have desktop GIS post their information to a county or state system on a regular basis. This local data needs to be augmented by regional data that also comes from the county or state system. This framework works well in that it keeps the responsibility of maintaining the data sets in the hands of those who are legally responsible for their accuracy and currency, but the problem of seamlessly integrating these disparate sources of data for the end user becomes more complicated.

From the user's point of view, GIS integration should be transparent. If the user wants to see the wetlands in town X, they should simply ask the system to display wetlands. The system should then have the intelligence to choose the appropriate wetlands data set to use. For example, the system might first check for the existence of a town X GIS server, and if it exists check it for a wetlands data set. If the server or the data set does not exist, the system would drop back to the next best available data, either from a county, regional or state agency. There could even be support for commercial data servers in this scenario.

Some early prototypes of GIS data servers already exist. Etak Corporation (http://www.etak.com) has implemented a geocoding service that can either run as a standalone application or as an integrated component of ArcView and MIT (http://ortho.mit.edu) has developed an ArcView extension that can make requests from a number of servers to add orthophotographs to a map. These GIS servers/services conform to the model proposed here in that they deliver geographic information over the Internet to mapping clients, but they are not very "intelligent," as they only connect to one specific data server and each user must find and install each individual software component.

A better system would separate the GIS client from the software that makes requests for services from GIS servers. This type of software is generally referred to as "middleware." In Figure 3 this middleware is represented by the red circle. All client requests go to the middleware component, or GIS agent. It is similar to a server in that clients direct their GIS requests to it, but it does not fulfill any requests on its own. Rather it is the repository for all the information regarding what GIS services are available. It handles all requests for GIS services, calls upon the GIS servers for data or analyses, and returns data to the client in the form of a map, report, chart, etc.

This type of middleware is sometimes called an intelligent agent, and many industries see agents as the future of the Internet. A Web search engine is an early example of a successful implementation of agent software. We, as Web users do not have to know about all the Web sites that contain information of interest to us. When we have a question we simply go to our favorite search engine, such as Altavista, type in our query, and get back a list of Web sites that are likely to interest us. In the GIS field this is equivalent to having an agent discover all of the databases of interest. Altavista also provides additional software that translates Web sites into a number of different languages. This ability to postprocess data also is important in GIS, because our data may need to be reprojected into a different coordinate system, or we may want an analysis algorithm performed on the data before we receive it. I believe the first generation of Internet-based GIS services will simply allow data to be gathered, reprojected and displayed in the form of a map. These systems will realize their full potential when a host of analytic GIS services are in place to complement the data gathering services.

While much of this discussion has been theoretical, I have tried to show that the evolution of Web-based mapping leads logically to this GIS service-focussed architecture. I have also tried to give examples of existing technology that show we are not far from being able to implement these systems. The most formidable barrier left is the reluctance of agencies to cooperate on this kind of project. The idea of intelligent agents has been around for years, but it has failed commercially because no one found a profit-based motive that would encourage the Web server industry to support the agent model. Hopefully, in the public sector the profit, or cost recovery motive will be less of an issue and we can build these systems on the belief that GIS users will benefit greatly from their existence.